Blast from the Past: ASDC Curation for NASA Suborbital Legacy Missions to Promote Discovery and Accessibility

Megan Buzanowicz[1,3], Sean Leavor[1,3], Nathan Jester[1,2], Gabriel Mojica[1,3], Abraham Porter[1,3], Gao Chen[3], John Kusterer[3]

[1]ADNET Systems, Inc.; [2]Booz Allen Hamilton; [3]NASA Langley Research Center (LaRC)

Atmospheric Science Data Center (ASDC)

- Established in 1991 at NASA LaRC
- One of twelve DAACs in NASA's Earth Observing System Data and Information System (EOSDIS)
- Focus in four subject areas:
  - Radiation Budget
  - Aerosols
  - Clouds
  - Tropospheric Chemistry
- Supports over 60 projects (orbiting and suborbital) and provides access to more than 1,000 archived collections.

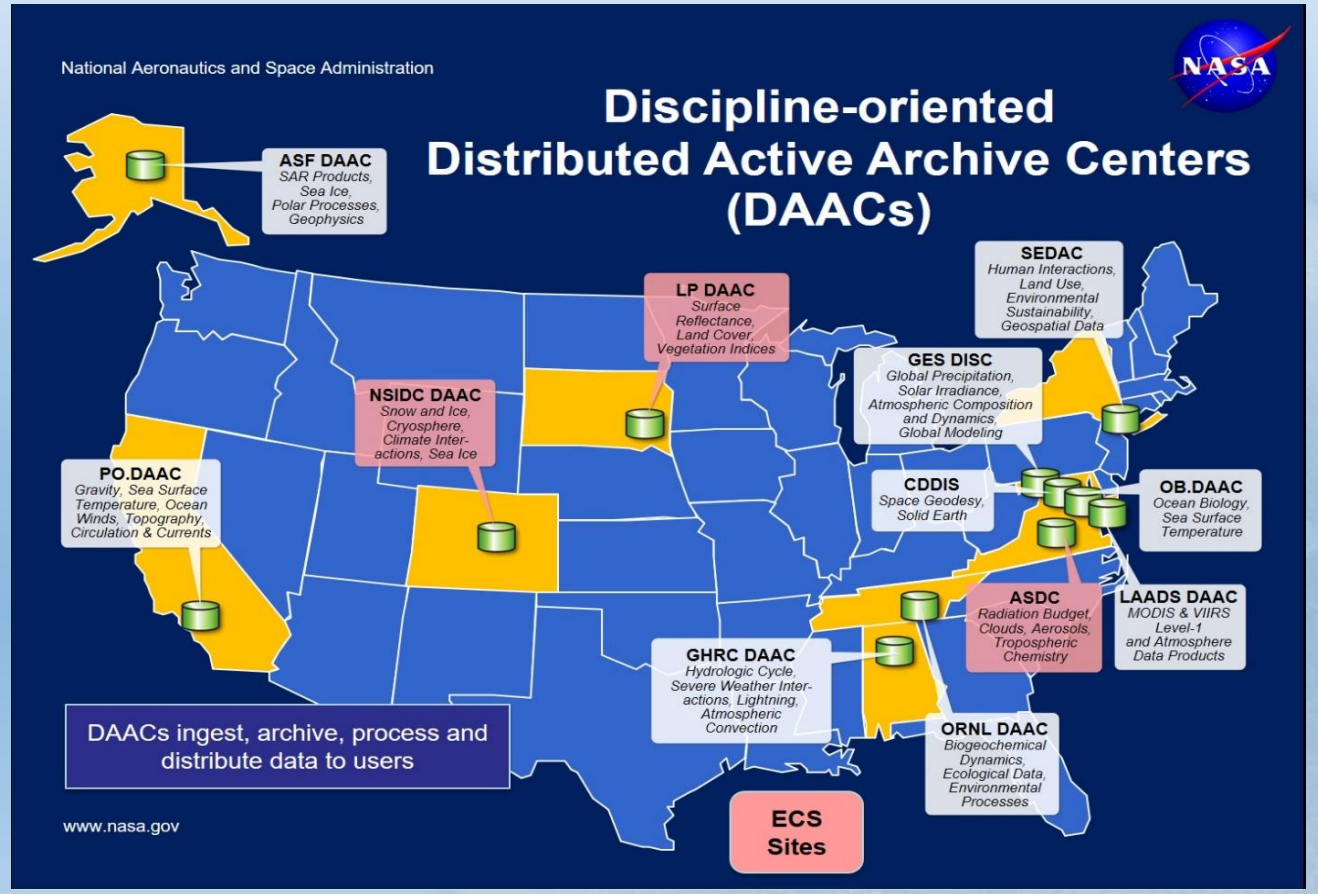

Image Credit: NASA Earthdata

Suborbital Support at the ASDC

- Field Investigations (Field Campaigns): Observational studies during which individuals, programs, agencies, and/or institutions utilize pre-selected, specific sensors or sets of sensors to acquire targeted observations in support of common, clearly defined, science or research objectives
- NASA has an extensive history of conducting such missions, beginning with the Chemical Instrument Test and Evaluation (CITE)
- Due to the high quantity of missions assigned to the ASDC, the Suborbital Support (SUBS) team was created

Ingest and Curation Overview

Defining Collections (Data Products):

- Developed a process to accommodate wide range of instrumentation (in-situ and remote sensing)
- Define collections based on measurement theme; consultation with science teams
  - E.g., trace gases or aerosols and measurement type (in-situ or remote sensing)
  - Established at times to support manuscript publication
- Collections are used for distribution on Earthdata Search

Metadata Extraction for Data Files:

- Most complex – but also most important step
  - Reflects the lack of rigorous data format requirements and metadata standards
- Team has developed a Python-based metadata extraction pipeline for automated metadata extraction from data files (stick around for a more detailed presentation on this!)
  - Can retrieve over 80% of the required metadata
- Often contact science teams to assist with missing metadata and ambiguities
- Process ensures high quality metadata at both the collection-level and granule-level (and feeds information back to the science teams)

DOIs and Landing Pages:

- Project-level and collection-level DOIs
- Issue additional DOIs to support manuscript development
- Reserve DOIs after preliminary data is submitted – enables the science team to include them in publication quality data
  - Supports data FAIRness
- Coordinate with science teams to ensure data product description accuracy and highlight noteworthy features

Distribution:

- ASDC Website
- Earthdata Search: collection-based search and discovery
- Sub-Orbital Order Tool (SOOT): variable-centric search capabilities

Documentation – distributed by the ASDC:

- Can vary widely and be unstructured
- Often contains science team generated reports, flight plans, contextual information regarding the fight and ReadMe files – critical to data usability
- SUBS team generates user guides and outreach materials highlighting suborbital projects

Data Revisions:

- Developed a tool and mechanism to monitor data submissions at the field repository for any post-project data revisions
  - Leverages the data integrity checking capability at the field data repository and ensures timely archival at the DAAC

Legacy Missions and Future Plans

- ASDC is actively working through a list of assigned legacy missions
  - Primary goal: Provide access to the datasets as they are currently formatted
  - Missions that are publicly available:
    - POLARIS (1997)
    - SOLVE (1999-2000)
    - SONEX (1997)
    - TOTE-VOTE (1995-1996)
    - STRAT (1995-1996)
  - Datasets span a wide-variety of file formats (text, Ames, GTE, ICARTT)
    - Data reporting standards have evolved significantly since 1983

- To enhance the **F**indability and **A**ccessibility (**FA**IR principles) of the data, the ASDC has begun researching ways to make them more accessible to a wider user community:
  - Distributing value-added products, including converting the files to a consistent and modern/standardized file format
  - Creating outreach materials and supplying user guides/ReadMe files
  - Incorporating standard names from the Atmospheric Composition Variable Standard Names Convention (ACVSNC)
  - Continue to improve the ASDC suborbital ingest and curation processes as more campaigns are archived
  - Continue to work with science teams and provide guidance on making data more FAIR
  - Assist with the improve and implementation of metadata standards, such as UMM-Var (Unified Metadata Model for Variables)

- Additional legacy campaigns to be archived at the ASDC in the near future:
  - GTE (TRACE, PEM, CITE, and ABLE)
  - ASHOE/MAESA
  - SPADE
  - AASE
  - STEP
  - AAOE

Sub-Orbital Order Tool (SOOT)

- Designed to address the distinctive characteristics of suborbital data while supporting the discovery, access, and use of data from campaigns archived at the ASDC
- Power User Interface (UI): intended for experienced data users and science teams with extensive knowledge of suborbital data
  - Allows users to refine their search based on campaign, principal investigator, start date, and dataIDs (including the variable in each dataID)
- Merge Service: enables users to align data to a common time base and view their selected data products from the same campaign, platform, date, and launch/flight in one file

Airborne and Field Data Workshop

- ADMG, ESDIS, and DAAC Airborne and Field Data Workshop
- April 23-24, 2024 (virtual)

Acknowledgements

- Michael Shook, NASA LaRC
- Morgan Silverman, Analytical Mechanics Associates (AMA)/NASA LaRC

Questions?

Resources

- Earthdata Forum
- Airborne and Field Data Resource Center
- Earthdata Search
- NASA Airborne Science Program
- ICARTT File Format Standards V2.0
- ADMG Terms and Definitions